

In the Claims:

1. A method for identifying designable protein backbone configurations comprising:

- 5 a. specifying a fixed number of amino acid secondary structural elements;
- b. generating a set of stacks comprising said secondary structural elements; and
- 10 c. evaluating designability of each stack within said set of stacks.

2. The method of claim 1, wherein said secondary structural elements comprise at least one alpha helix, at least one beta strand or both.

15 3. The method of claim 1, wherein one secondary structural element corresponds to an alpha helix.

20 4. The method of claim 1, wherein one secondary structural element corresponds to beta strand.

5. The method of claim 1, wherein said fixed number of secondary structural elements is one to twenty.

25 6. The method of claim 5, wherein said fixed number of secondary structural elements is four.

7. The method of claim 2, wherein said alpha helix is about 15 amino acids in length.

8. The method of claim 1, wherein a center of mass and an Euler angle are randomly selected for each element of said stack.

5 9. The method of claim 1, wherein step (b) includes generating an initial stack by a conjugate gradient method.

10 10. The method of claim 9, wherein the conjugate gradient method includes a step of determining the minimum packing energy of said stack.

15 11. The method of claim 9, further including a step of generating additional stacks by performing one or more symmetry operations.

12. The method of claim 11, wherein said symmetry operations comprise slide operations or screw operations.

20 13. The method of claim 1, wherein step (b) further includes a step of confirming that said stack does not exceed a predetermined constraint wherein a stack that exceeds said predetermined constraint is discarded.

25 14. The method of claim 13, wherein said predetermined constraint is an end-to-end distance between connected helices.

15. The method of claim 1, wherein step (b) further includes a step of determining the surface exposure of each amino acid within each stack to water.

16. The method of claim 9, wherein a plurality of stacks are generated, wherein each stack is based on a distinct set of randomly selected starting coordinates.

5

17. The method of claim 16, wherein said randomly selected starting coordinates include a center of mass and Euler angles for each element of said stack.

10

18. The method of claim 9, wherein further including a step of assessing the completeness of said plurality of generated stacks.

15

19. The method of claim 18, wherein plurality of generated stacks is complete when about 90% to about 95% of newly generated stacks lie within a root-mean-square distance of about 1.5 Angstroms of at least one stack already in the set.

20

20. The method of claim 1, further comprising a step of grouping said set of stacks generated in step (b) into clusters.

25

21. The method of claim 19, wherein said clustered stacks are sorted and listed according to total surface exposure to water from the most compact stack to the least compact stack.

22. The method of claim 21, wherein all stacks that are within 1.5 Angstroms crms of said most compact stack are eliminated from said cluster and wherein said process is repeated for a next most compact stack on said list until the end of said list is reached.

23. The method of claim 1, wherein a random set of amino acid sequences is generated based on binary sequences consisting of Hydrophobic (H) and Polar (P) amino acids wherein a random sequence of amino acids has a length of 2^n wherein $n=1-500$.

5

24. The method of claim 23, wherein each amino acid sequence is reduced to the hydrophobicities of its individual amino acids.

25. The method of claim 21, wherein each amino acid in each stack has a surface exposure value.

10

26. The method of claim 23, wherein the energy of an amino acid sequence folded into a particular configuration is

15

$$E_{\text{designability}} = - \sum_i h_i s_i,$$

where h_i is the hydrophobicity of the i th element of the sequence and s_i is the surface exposure of the i th amino-acid sphere in the particular stack.

20

27. The method of claim 26, wherein for each random amino acid sequence considered, the stack with the lowest energy is a designable structure.

25

28. The method of claim 26, wherein a highly designable stack is identified when the number of amino acid sequences with said stack as the lowest energy state, is larger than the average number of sequences per stack.